United Nations ECE/CES/2015/42

Economic and Social Council

Distr.: General
7 May 2015

English only

Economic Commission for Europe

Conference of European Statisticians

Sixty-third plenary session
Geneva, 15-17 June 2015
Item 4 of the provisional agenda
Modernisation of statistical production and services and managing for efficiency

Innovative products and services for European statistics: challenges and opportunities

Note by Eurostat

Summary

This paper focuses on innovative products and services that are a key aspect of the modernisation of European statistics. Accelerating demand for information by different stakeholders is putting pressures on statistical organizations around the world. They need to deliver timelier, high-quality information that responds to varying user needs and supports evidence-based decision making. Consumers are increasingly empowered by advanced technologies that leverage the value added of data integration and analytics for knowledge generation. To respond to this fast changing demand from users, statistical organizations need to strengthen the portfolio of products they offer by adding value to the statistics they disseminate.

This paper is presented for discussion to the second session of the Conference of European Statisticians' seminar "Modernisation of statistical production and services and managing for efficiency".

GE.15-09142 (E)

1509142

Please recycle

I. Introduction

1. In a digital and information driven society, official statistics face many challenges in responding to user's needs relating to technology, methodology, governance, organization of work, dissemination and innovation. Currently, national statistical offices and international statistical organizations are striving to adapt to the evolving society to stay relevant by becoming more modern.

2. The provision of innovative products and services is a key aspect of the modernisation of European statistics. The key drivers are: a progressive switch from statistics to dashboards of indicators produced by combining different sources and methods in a flexible and agile way; a renewed emphasis on timeliness combined with quality assurance; a value added approach to statistics moving towards the provision of statistical information and services; an interconnected view that encompasses basic statistics, accounting frameworks and indicators; products that are not anymore mere numbers but combine also metadata and visualization.

3. The European Statistical System (ESS) is moving along these new societal trends: until recently rapid estimates, composite indicators, scoreboards, wiki articles and mobile services represented the innovative strand in official statistics. This paper explores some potentially new ideas for renewed products and services of European statistics: portals, interlinked indicators, partnership-based outputs, classic statistics produced with innovative techniques, infographics, tailor-made statistics, linked open data (LOD) based dissemination and data analytics services.

4. For official statisticians, this is a new way of conceiving, producing and disseminating statistical information that reflects the technological and communication-oriented daily reality of societies.

5. These technology trends, coupled with the increased demand from users, are a great opportunity for statistical organizations to enhance their relevance as data are becoming a key source of knowledge, the new capital. Statistical offices will succeed in achieving this goal only if they are available to invest in strengthening their product and service portfolio, putting users in the driver seat and in recognising that quality is the overarching principle to ensure that products and services are valuable to society.

II. Innovative products and services

6. In the aftermath of the financial and economic crisis, official statisticians underwent an in-depth reflection on how to respond in an agile, flexible and tailored way to the increasing request for statistical information. Such an analysis targeted different dimensions of collection, compilation and dissemination of official statistics with a two-fold aim: First, to streamline the production of official statistics by optimising the production processes and, second, enhance the capacity of national and international statistical organizations in disseminating and communicating statistical products and services. The digital transformation of the society, enabled by the data revolution, is highly influencing these developments.

7. The following section will address the common features that can be identified in the different modernisation programmes that statistical offices have launched in the recent years, both as concerns statistical production and statistical output (products and services).

A. Recent trends

8. As regards the dissemination and communication of the output of statistical production, general trends could be highlighted.

1. Statistical data for policy making: indicators vs. dashboards vs. synthetic/composite indicators

9. The demand of indicators for policy making purposes has considerably increased during the last years. Evidence-based policy making has become a "must" in running economic, social and environmental policies at global, international, national and regional level. In a user-producer-dialogue, policies are more and more defined in "measurable" terms and associated to a set of indicators selected or created to provide the necessary information on the achievement of the policy targets. In addition, the achievement of policy targets is often based on thresholds derived from a statistical analysis and assessment of past performance.

10. If until few years ago, single key indicators were traditionally used for these purposes, nowadays the following types of indicator groups are developed to ensure a proper follow-up and measurement of efficacy and success of policies:

- Scoreboards (set of indicators on the situation);
- Dashboards (set of indicators for monitoring and control);
- Indicators catering for sub-populations (to measure distributive impacts – e.g. inflation by household spending group to measure distributive impacts of inflation);
- Composite indicators (compilation of individual indicators into a single index, on the basis of an underlying model) and;
- Synthetic indicators (to condense the measurement of several sub-indicators into one value).

11. This tendency appears quite strongly in the instruments used by the European Commission to run and monitor the European policies. For example the European Semester relies on a scoreboard of 11 headline indicators measuring internal and external macroeconomic imbalances and competitiveness of EU countries. It is monitored with 28 auxiliary indicators. The Europe 2020 strategy, on the other hand, aims at achieving 5 high level targets and measures this on the basis of 9 headline indicators translated into national targets and indicators. Similarly, other key European policies such as the Single Market Monitoring, the Digital Agenda, etc. are measured by statistical indicators. For all these instruments, Eurostat provides the corresponding dashboard and scoreboard indicators in a user friendly and policy oriented fashion, combining data and metadata and offering, in specific cases, quality indications on the indicators themselves.

12. Whilst traditionally indicator sets are quite well established in statistical offices, with successful examples such as those published by Eurostat (Principal European Economic Indicators, Europe 2020 indicators, GDP and beyond indicators, Macroeconomic Imbalance Procedure indicators), users have recently started to value the following new characteristics:

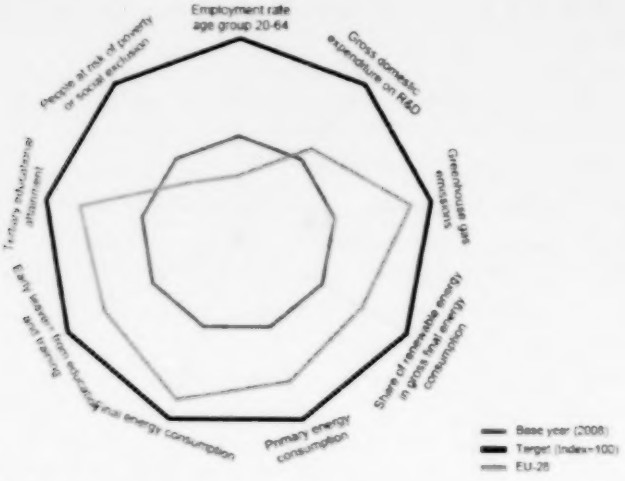
- Inclusiveness;
- Interlinks among different dimensions of the policy to be monitored;
- Agility and flexibility to respond to changing requirements;
- Attractiveness for policy makers (quality, including timeliness and accuracy; possibility to derive new indicators in a quicker and targeted manner);

- Easy accessibility;
- Synthesis.

13. These new characteristics call for agile production systems based on a smart and optimised use of the available statistical information and able to amplify the synergies in deriving "new" indicators from the existing set of basic statistics.

14. The challenge for statistical offices in this context is to set up production and information systems that will shift from the traditional paradigm of offering "statistics" to a more service oriented attitude "to connect, aggregate and tailor" statistical information, on the basis of a data warehouse approach and providing services in addition to data.

Figure 1
Examples of indicators disseminated by Eurostat


Employment rate age group 20-64
Gross domestic expenditure on R&D
Greenhouse gas emissions
Share of renewable energy in gross final energy consumption
Primary energy consumption
Final energy consumption
Early leavers from education and training
Tertiary educational attainment
People at risk of poverty or social exclusion
Base year (2008)
Target (Index=100)
EU-28

Europe 2020 indicators

GDP and Beyond indicators

Source: Eurostat

2. Rapid estimates

15. Timeliness has a prominent role among the quality dimensions associated to statistics and statistical indicators. The general perception about the timeliness aspect in the release of official statistics is quickly evolving in line with the trend of a more data driven society that relies on access to "numbers" on the fly. Indeed, if 25 years ago the switch, for example, from annual to quarterly national accounts took some time to be fully endorsed by users and producers, today fresh and quick statistical information is continuously required.

16. At the same time, new actors, mainly from the private sector, are starting to use new methodologies, techniques and data sources to produce proxies of official statistics. Concepts such as nowcasting and flash estimates are regularly used in statistical offices, and flash estimates of GDP and the Harmonised Index of Consumer Prices (HICP) are released by Eurostat for the European Union (EU) and euro area together with the available figures from EU member states. Rapid estimates and nowcasting are particularly important in the current economic and social conditions characterised by high volatility since they can provide fast feedback to policy makers on the policy impacts ("did we get it right or not?").

17. The challenge for the official statistics community is to ensure an adequate level of quality of the statistical indicators provided in a timely manner, as the quality may have consequences on the public image of and trust to statistical offices (a reflection linked to the idea of "labelling" of official statistics).

3. Modern dissemination and communication tools

18. Recently, particular emphasis has been put in statistical offices to adapting the style and services associated to dissemination and communication. The technological drivers, as well as the growing information demand, forced statistical offices to redesign the way of presenting and disseminating official statistics. Some offices have used professional communicators to focus their message and to help them adopt a modern style in communication. A number of statistical offices are showcasing new innovations: on one side by using new technological tools (interactive tools, mobile services, visualisation tools, etc.), and on the other side by applying a more targeted and modern communication approach (wiki-style dissemination sites, shorter and condensed presentations, almost disappearance of traditional publications, more visual-oriented presentation than traditional statistical tables, dynamic dissemination and data visualisation).

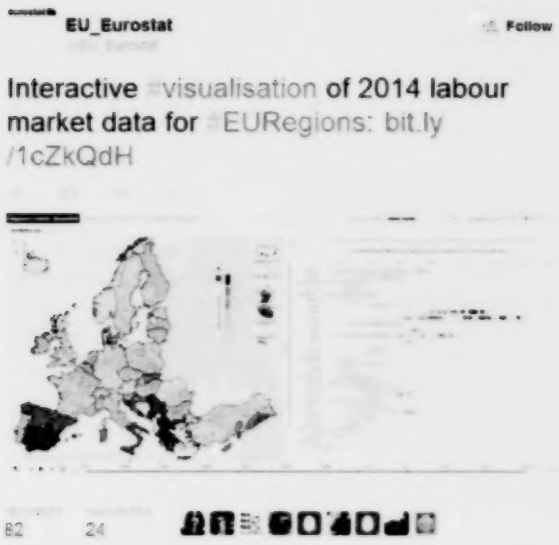
19. At the same time, the use of social networks for statistical dissemination (e.g. Twitter and Facebook) has brought a new dimension to disseminating statistics and in informing the general public with statistical news. Eventually, new profiles of data users appear, like the emergence of users of statistical data that create stories on the basis of data analytics services offered by statistical organizations (e.g. data journalists).

20. Eurostat is quite active in this context, for example

- Economic Trends offers a dynamic representation of a user-driven choice of the combination key indicator-country through a map visualisation;
- Statistics Illustrated provides a dynamic-over-the-time representation of key indicators (e.g. the MIP indicators), combining the evolution of the indicators by selected countries;
- Widgets give a quick information on key indicators (with the possibility of importing in its own web representation the widget facility);
- Inflation Dashboard provides a tool to display the Harmonised Consumer Price Index data simultaneously, by products and countries;

- Statistical Atlas offers an interactive map viewer for statistical and topographical maps;
- Mobile apps on the EU Economy and on Countries profile provides insights in European statistics through an easy access via mobile devices;
- Extraction tools - Tables, Graphs and Maps (TGM) offer a single style tool for downloads and prints of tables, graphs and maps;
- Data visualisation via Twitter addresses users via the social network;
- Infographics translate statistical information on key news and indicators in a visually oriented style.

Figure 2
Example of data visualisation tools provided by Eurostat



Source: Twitter.com/EU_Eurostat

Source: Eurostat Statistical Atlas

B. Future perspectives

21. Until now, official statisticians have quite successfully managed to innovate and develop new statistical products and services to meet new users' requirements. Nevertheless, the digital transformation and rapid changes that the society is experiencing require additional efforts. The following paragraphs briefly explore some potentially interesting ideas for the near future.

1. Portals

22. One of the key changes in the way of approaching statistical dissemination and communication in statistical offices is to offer tailor made services that generate added value to the traditional offer of statistics. In some countries, a tailor-made offer of statistical data and services targeting specific user groups, such as enterprises, has already started. The key characteristic of these services is to provide an attractive mix of dissemination facilities through a portal tailored to match the user interest: in case of enterprises, portals can offer specific tailor-made statistics for the interests of an enterprise or enterprises. The concept could be expanded to provide tailor-made services based on a data warehouse system that, in the ESS perspective, covers a system of statistical information repositories (data warehouses), is metadata driven and as much automated as possible. The Eurostat powered EU Census hub could be considered as a first realisation of such concept where users can define their queries and get comparable European information on census related indicators. Another example in this direction is the initiative "confidentiality on the fly" on which some statistical offices are currently working.

23. Most recently, portals increasingly offer diversified entry points for consumers and provide a guided tour that assists them in using the tool. In addition, portals provide a good data search and browsing capacity and are increasingly becoming search engines for statistics as opposed to a main websites of statistical offices.

24. The challenge for statistical offices is to set up the architecture and infrastructure required to successfully run portals that are user-friendly and serve a variety of data needs, as well as to switch from the traditional dissemination approach to a more tailored one.

2. Data integration

25. One of the major interests for users is getting information about the level of data integration. This information can be derived through metadata in single dissemination data warehouses. This approach could also serve the purpose of interlinking indicators, as a by-product.

26. Links between indicators describing different aspects of the same phenomenon are of increasing interest for users. This is particularly the case when analysing the impact of policies using a set of indicators.

27. The traditional narrow stovepipe approach usually applied by statistical offices in releasing statistical indicators has proven limited against the need for data integration. Measures of data integration and coherence are becoming a prerequisite for analysts. Therefore, statistical offices need to explore methodologies for better data integration to ensure, through the modernisation of statistical production, more integrated statistics (once again, a data warehouse organisation could facilitate this approach). In case of indicators, further steps along these lines are taken as the use of composite and synthetic indicators is becoming more widespread and benefits from interesting suggestions coming from other scientific disciplines (e.g. the use of POSET – partially ordered sets).

28. The challenge for statistical offices is to foster the development of methodological approaches to compile and describe sets of cross-domain indicators and to adequately represent and explain interlinks (in case on the basis of social and economic theories).

3. Partnership-based output

29. Statistics are not anymore exclusively produced by statistical authorities. The strong demand generated by a data-driven society attracts new major actors in the statistical market. Whilst we, official statisticians, could claim that the role of official statistics will continue to be relevant, notably as reference benchmark, we have to recognise that potential partnerships in collecting, processing and disseminating statistical products and services is an opportunity to be explored. Such partnerships could be useful for example in order to:

- Exploit the data revolution (in collaboration with big data collectors such as mobile phone providers or retail sale owners as regards their datasets);
- Benefit from powerful data dissemination partners (e.g. major web search engines or global vendors and their "data analytics" initiatives);
- Use privately produced information to derive statistics (e.g. real estate agencies data for compiling commercial property price indices).

30. For this purpose, Eurostat has successfully developed partnerships with, for example, a major web research engine for the dissemination of 10 sets of indicators (Google Public Data Explorer).

31. The challenges for statistical offices in this context vary according to the collaborative mechanisms that could be set up. Aspects like legal agreements, reputational effects and effective partnership have to be considered in this respect.

4. New data sources as generators of innovative products

32. New products and services could be developed using the available new sources, such as big data, administrative sources and geo-spatial information. New data sources can help generate new information: for example higher frequency information using big data (consumer sentiment on a weekly basis) or more geographical detail (territorial detailed labour force information for local policy making using administrative data combined with the labour force survey).

33. Eurostat is, with the ESS partners, currently exploring the potential offered by:

- New sources in particular by big data (e.g. tourism statistics based on mobile phone data and forecasts of key macroeconomic indicators using big data);
- The more intensive use of administrative data;
- Geo-referencing statistical data (e.g. the Inspire geo-portal, a EU Commission initiative).

34. The challenges for official statisticians in this context are mainly related to the resources to be invested in the methodology, technology and infrastructure as well as the required skills for the future generation of statisticians (data scientists).

5. Infographics, visualisation tools, media supports

35. Visual communication has become paramount in modern statistical communication and dissemination. Newspapers, media and presentations more and more make use of visual concepts, videos and new tools like infographics (visual representation of information). The opportunities offered by visual communication reflect the media-oriented attitude of the society as well as the intensive use of mobile applications and social networks. Whilst some statistical offices have started to disseminate official statistics using these tools, work still has to be done to keep the pace with the developments in the media.

36. In addition, new forms of communication, ranging from e-books to interactive press releases, focus on faster and more immediate communication between data and text, and build on data integration moving away from stovepipe or a topic by topic approach to dissemination. Social media press releases are a good example of this trend.

37. Eurosat is a front liner in exploring and deploying new visualisation tools that address these challenges: infographics for dedicated dissemination; videos to illustrate concepts and data; statistics explained (wiki dissemination of statistical information).

38. The challenge for statistical offices in this area is to be able to translate complex statistical concepts in attractive statistical information, on the basis of an appropriate choice of subjects, conception of synthesised messages combined with attractive representation and dissemination facilities.

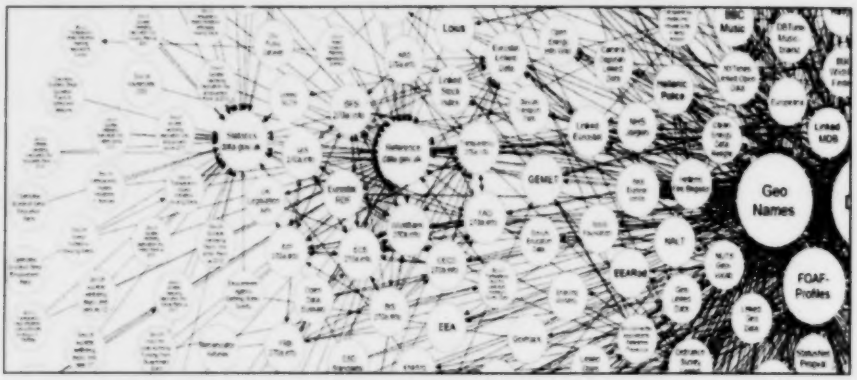
6. Linked Open Data (LOD) - semantic web

39. Linked Data is about using the Web to connect related data that wasn't previously linked, or using the Web to lower the barriers to linking data currently linked using other methods. More specifically, Linked Data is defined in Wikipedia as "a recommended best practice for exposing, sharing, and connecting pieces of data, information, and knowledge on the Semantic Web" using data and metadata standards.

40. From a statistical point of view the Linked Open Data (LOD) initiative builds on the general tendency of governmental organisations and statistical offices to provide open data (for example, Eurostat data represents 90% of the EU open data portal offer) and to apply to them a web semantic approach. The semantic web is an extension of the current web in which information is given well-defined meaning, better enabling computers and people to work in cooperation. Such an approach would foster the link among data sets and in the end, would create a larger statistical information-based virtual data storage that could be exploited for statistical purposes, also according to a machine learning approach. In addition, LOD would favour the potential to generate reuse of data by semantic integration, catering to different users also through secondary dissemination platforms.

41. The challenge for statistical offices will be mainly in running the standardisation wave and couple it with the web opportunities.

Figure 3
Linked Open Data



Extract from the Linking Open Data cloud diagram 2014, by Max Schmachtenberg, Christian Bizer, Anja Jentzsch and Richard Cyganiak. http://lod-cloud.net/

7. Machine learning

42. Technical advancement, such as cloud services and computing, could facilitate a statistical use of data available on the web. Recently, a big web retailer (Amazon) launched a new cloud service to help users build predictions on the basis of model tools that use data mining techniques. Users can use the service for a price and harness their own data sources. This is related to the move towards a decision system support in a cloud environment and it is close to the concept of data analytics as a service that is becoming very relevant for official statistics.

43. The challenge for official statisticians is to create a cloud platform where data are stored and analysed in an integrated way via data mining models that offer description and prediction tailored to users (e.g. client profiling, behaviour responses). A further step towards a service oriented approach for statistical offices.

8. Data analytics services

44. Statistical offices are becoming statistical service providers. "Statistics as a service" will represent an increasingly important development along with cloud environments for data integration and data analytics. The "platform as a service" type of framework will be increasingly introduced in the statistical community, in line with the expectation that processing efficiency will increase more rapidly than data transfer capacity (e.g. data gravity models). This leads to the creation of large aggregate data centres on cloud platforms. Data analytics combined with web semantics represents the future research development in this area.

45. The challenge for official statisticians is to address these developments from a statistical point of view, with consideration to privacy, confidentiality, legal aspects, analytical developments, provision and use by data owners and data analytics platforms for analysis and prediction as well as building capacity for data analytics.

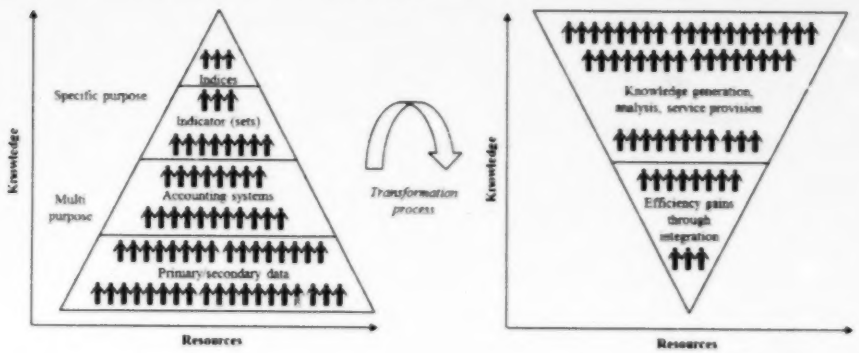
III. Issues and challenges

46. The introduction of innovative products and services in the official statistics community generates specific challenges. The main challenge is creating working environments that are oriented to creativity and innovation, support the development of new products and enhance the skills of staff to perform tasks from data science to communication in addition to new and consolidated methodological skills.

47. Another key challenge is to make the statistical organisation evolve and successfully undertake a paradigm shift from the traditional way of producing statistics to a more service oriented culture based on modern methods and techniques. The challenge will be to move from the current system towards a knowledge-oriented statistical service, emphasising the orientation towards the clients, i.e. to serve the users in the best possible way. This implies a shift in the focus of statistical offices from an internal perspective (production system) to an external one (users). Organisationally speaking, this paradigm shift relies on strengthening the services layer through a gradual transformation process targeted to generate knowledge, analysis and services.

Figure 3

Transformation process in statistical organisations



IV. Conclusions

48. In order to address these challenges from a broader perspective to modernisation, ESS undertook a major initiative to meet users' requirements and face the methodological and technological challenges to develop its products and services towards the ESS Vision 2020. The ESS Vision 2020 is the framework for the modernisation of the ESS through five strategic goals relating to users, new sources, quality, production processes and dissemination/communication. This Vision is currently being implemented through a portfolio of enabling projects to build the ESS capabilities for achieving the Vision.

49. One of the key strands for modernisation in the ESS focuses on digital communication, users, analytics and innovative products. This programme will facilitate the move from a traditional mix of statistics to innovative products and services based on a user-driven, technologically advanced and service-oriented culture, supported by adequate infrastructure and architecture. This strand, combined with new data sources and quality framework, is planned to bring innovation into ESS products and services. The achievement of the ESS Vision 2020 depends on the ability of the ESS to be open to innovation and successfully apply a change management in a collaborative effort.